استراتيجية تطورية ذات طفرات وراثية مترابطة لحل مشكلة سرج روزنبروك

د.ميادة فيصل عبد الحليم*

المستخلص

مشكلة سرج روزنبروك من المشاكل احادية الحل الامثل. يقع الحل الامثل على امتداد وادي منحني وضيق الذي يتناقص بشكل ضئيل الى ان يصل الى قاع منبسط. هذا الوادي يقع على اطراف فضاء البحث. من السهولة الوصول الى هذا الوادي ولكن الوصول الى الحل الامثل صعب. يرجع سبب الصعوبة في الاقتراب من الحل الامثل لكون معاملات المشكلة تمتلك صفة تلرابط متداخل قوي. من الممكن استخدام الأستراتيجيات التطورية متعددة الأفراد والتي يرمز لها $(\mu \overset{+}{,} \lambda)$-ستراتيجية تطورية للوصول الى الحل الامثل. هذا البحث يوضح مقدار معاملات الاستراتيجيات التطورية المناسبة لحل مشكلة سرج روزنبروك بشكل موثوق به.

*جامعة بغداد – كلية العلوم –قسم علوم الحاسبات

Optimization. Evolutionary Computation, Vol.1, No.1, Spring 1993.

[4] BӧCk, T., Hoffmeister, F., And Schwefel, H.-P. (1991) : A Survey Of Evolution Strategies. Proceedings Of The 4th International Conference On Genetic Algorithms, Pp. 2-9, La Jolla, CA: Morgan Kaufmann.

[5] Kappler, C. (1996): Are Evolutionary Algorithms Improved By Large Mutations?, Parallel Problem Solving From Nature-PPSN III, Eds. H. Voigt, W. Ebeling, I. Rechenberg And H. Schwefel, Berlin, Springer-Verlage, September, Pp.346-355.

[6] Schwefel, H. -P. And BӧCk, T. (1997): Artificial Evolution: How And Why. In D. Qualiarella, Periaux, C. Poloni, And G. Winter (Eds.), Genetic Algorithms And Evolution Strategies In Engineering And Computer Science. John Wiley And Sons Ltd. Pp. 1-19.

[7] Sullivan, J. And Pipe A. (1996): An Evolution Strategy For On-Line Optimization Of Dynamic Objective Functions, Parallel Problem Solving From Nature-PPSN III, Eds. H. Voigt, W. Ebeling, I. Rechenberg And H. Schwefel, Berlin, Springer-Verlage, September, Pp.781-790.

[8] Lazauskas, L. (1998): Test Functions And Benchmarks, Http://Titan.Princeton.Edu/Testproblems, Available At The Internet.

[9] Ashlock, D. (2001): Optimizing Real Valued Functions, Http://Www.Math.Isatate.Edu/Danwell/Ma378, Available At The Internet.

also beneficial such that provide the algorithm with more chances to exchange evolved angles between individuals.

5.Conclusion

The multimembered $(\mu \overset{+}{,} \lambda)$-ESs provides the user with the ability to vary the amount of strategy parameters, i.e., standard deviations and inclination angles attached to the ES's individuals. Moreover, the concept of self adaptation of these strategy parameters by means of extending the evolutionary operators mutation and recombination to the step sizes and rotation angles provide the robustness and learning capabilities of the mechanism. For Rosenbrock's Saddle problem, the mutative step size control as presented by ES1 is insufficient for the complicated case of the problem, because the internal models of individuals are to restricted by $n_\sigma = n$ to explore the 4-dimensional search space *in an effective* way. *On the other hand, adapting* a relatively large amount of strategy parameters including rotation angles, as implied by ES2 and ES3, were found here to be sufficient to locate the reliable solutions. Additionally, ES3 requires more computational work than ES2, as it uses more strategy parameters, but it finally pointed out to be more effective then the latter ES.

Reference

[1] Schwefel, H.-P. (1981): Numerical Optimization Of Computer Models, Chichester: Wiley.

[2] B;Ck, T. And Schwefel, H. -P. (1995): Evolution Strategies I: Variants And Their Computational Implementation, Genetic Algorithms In Engineering And Computer Science Pp. 11-126.

[3] B;Ck, T. And Schwefel, H. -P. (1993): An Overview Of Evolutionary Algorithms For Parameter

300	8.3371e-003	4.4710e-010	2.7581e-004
325	6.1316e-003	3.6148e-012	1.9108e-004
350	4.6242e-003	6.5276e-013	1.3445e-004
375	3.4573e-003	3.4632e-013	7.1630e-005
400	2.6312e-003	2.0917e-013	3.9637e-005
425	1.9939e-003	2.1868e-013	8.4606e-007
450	1.5239e-003	1.5640e-013	2.6221e-011
475	1.1368e-003	2.9217e-013	2.2194e-014
500	8.3619e-004	3.0444e-013	9.9045e-015

From the results above, one can see the difference in the reliability behavior of ES1 at one hand and ES2 and ES3 at the other hand. The performance of an evolution strategy like ES1 becomes worse when only standard mutation with no orientation ability is used. In such cases, the step length changes provided by the standard deviations occurred to an evolved individual did not let that individual to locate the promising region of the search space. While the results of ES2 and ES3, show that the use of correlated mutation with an evolution strategy for this type of problem will causes more effective search, i.e., better approximation of the minimum and more robust search. In other wards, the orientation feature permits the evolution strategy to change its direction so as to locate the reliable region of the search space. Here, while the standard deviations vector remove the evolved individual from one position to another via addition or subtraction of step lengths, the inclination angle vector enable that individual to change its orientation. Moreover, from the results of both these two evolution strategies, one can deduce that the use of recombination operator to recombine parents' angles, rather than standard no recombination operation applied to inclination angles, is

3. **ES3: (30+200)-ES with number of parents $\mu = 30$ and number of offspring $\lambda = 200$. Use self adaptation of $n_\sigma = n$ standard deviations, and the standard deviations $\sigma_i \quad \forall i \in \{1,...,n\}$ are initialized to a value of 0.5. Use correlated mutation with $n_\alpha = n.(n-1)/2$. Also, as in ES2 the standard recombination setting are used for only objective variables x_i and standard deviations σ_i, while we used here global discrete recombination on inclination angles $\alpha_j \forall j \in \{1,...,n.(n-1)/2\}$.**

All results were obtained by running fifteen experiments per evolution strategy and averaging the resulting data. Each algorithm run performed 500 generations and at each generation we store the minimum objective function value of the current evolved population values. Then, the results in the following table are average of minimum of the different fifteen runs listed at every twenty-five generation.

Table 1: Experimental Results

Gen. Num.	ES1	ES2	ES3
1	1.0677e+002	1.0677e+002	1.1374e+002
25	8.3806e-001	1.1771e+002	1.4471e+000
50	3.2190e-001	2.6649e-001	1.1921e-001
75	1.8906e-001	1.7441e-001	5.7784e-002
100	1.2456e-001	1.0936e-001	2.5961e-002
125	8.4010e-002	6.1457e-002	1.2126e-002
150	5.9023e-002	4.0747e-002	6.1171e-003
175	4.1338e-002	2.5834e-002	2.6805e-003
200	2.9641e-002	1.7933e-002	1.3899e-003
225	2.1335e-002	9.5590e-003	9.0278e-004
250	1.5286e-002	1.9512e-003	6.5137e-004
275	1.1232e-002	1.5993e-006	4.0483e-004

The harder problem of Rosenbrock's Saddle is with $n = 4$ [8][9] which has been attempted by some researchers [8]. However, their evolutionary algorithm (Breeder GA) took roughly 77.000 function evaluations to reach $f(\vec{x}) = 0.1$. Hence, in the next section we present evaluation strategies with our proposed setting.

4. Experimental Results

Having presented the concepts of the multimembered evolution strategies with correlated mutation together with the formulation of the mathematical modal of Rosenbrock's Saddle problem in the previous sections, we have in this section to look at the experimental results of three evolution strategies with different settings. The evolution strategies used in the experiments are listed below:

1. **ES1:** (30+200)-ES with number of parents $\mu = 30$ and number of offspring $\lambda = 200$. Use self-adaptation of $n_\sigma = n$ standard deviations, and the standard deviations $\sigma_i \quad \forall i \in \{1,...,n\}$ are initialized to a value of 0.5. No correlated mutation, i.e., $n_\alpha = 0$. Finally, the standard recombination setting suggested by Schwefel is used.
2. **ES2:** (30+200)-ES with number of parents $\mu = 30$ and number of offspring $\lambda = 200$. Use self adaptation of $n_\sigma = 2$ standard deviations, and the standard deviations $\sigma_i \quad \forall i \in \{1,...,n\}$ are initialized to a value of 0.5. Use correlated mutation with $n_\alpha = n-1$. Finally, the standard recombination setting are used for only objective variables x_i and standard deviations σ_i ,while we used here global discrete recombination on inclination angles $\alpha_j \forall j \in \{1,...,n-1\}$.

One last thing remains to be explained is the following. Schwefel's claim is that there is an implicit link between an appropriate internal model and good fitness values and so good values of the strategy parameters will emerge from the population and accelerate the rate of convergence. He likens this linear correlation process to the epigenetic apparatus or transmission mechanism between genotype and phenotype, whereby a single gene can influence several phenotypic characteristics (pleitropy) and vice versa (polygeny)[7].

1.Rosenbrock's Saddle Problem

The Rosenbrock's Saddle problem, also called Rosenbrock's valley or Banana problem, as formulated in equation 10, is a unimodal optimization problem that has the minimum=0.0 located at $x_i = 1.0,\ 1 \le i \le n$ within a long narrow curved valley which is only slightly decreasing and with a flat bottom. This valley is located near the boundaries of the search space. To find the valley is trivial; however, convergence to the global minimum is difficult because the problem has a strong inter-parameter linkage or non-separable characteristic, i.e., there are nonlinear interactions between the variables of the problem. This characteristic has been considered by some authors to be real difficult and challenge one for any continuous optimization program, and the feature will probably cause slow convergence in many optimization algorithms since they must permanently change their search direction to reach the minimum.

$$f(\vec{x}) = \sum_{i=1}^{n-1} 100(x_i^2 - x_{i+1})^2 + (1 - x_i)^2 \quad (10)$$

where $-2.048 \le x_i \le 2.048$

which are unit matrices except that $t_{pp} = t_{qq} = \cos(\alpha'_j)$ and $t_{pq} = -t_{qp} = -\sin(\alpha'_j)$,i.e., the trigonometric terms are located in column p and q each.

The factors τ_0, τ , and β are rather exogenous parameters, which Schwefel suggested to set them as follows

$$\tau_0 = \frac{1}{\sqrt{2n}}, \quad \tau = \frac{1}{\sqrt{2\sqrt{n}}}, \text{ and } \beta \approx 0.0873 \ [\approx 5^0].$$

The conceptual algorithm of the mutation operation on strategy parameter $\vec{\alpha}$ and is embedded below (see figure 2) in the framework of the above sequences.

```
Algorithm 1: Coordinate Transformation of Positional Angles α→
n_lower = n − n_σ + 1;
n_upper = n − 1;
n_q = n_α;
for I= n_lower to n_upper
   n1 = n − I;
   n2 = n;
   for J=1 to I
      x1 = x(n1);                          x2 = x(n2);
      y1 = sin(α(n_q));                    y2 = cos(α(n_q));
      x(n2) = x1 * y1 + x2 * y2;           x(n1) = x1 * y2 − x2 * y1;
      n2 = n2 − 1;
      n_q = n_q − 1;
   endfor
endfor
```

Figure 2: Conceptual Algorithm of Coordinate Transformation

where $(cor_1(\vec{\sigma}', \vec{\alpha}'),\dots,cor_n(\vec{\sigma}', \vec{\alpha}'))$ $\vec{cor}$ = is a random vector with normally distributed, eventually correlated components. The vector $\vec{cor}$ can be calculated according to [6] $\vec{cor} = T'\vec{Z}$ where:

$\vec{Z} = (z_1,\dots,z_n)$ with $z_i \sim N(0,\sigma'^{2}_{j})\forall i \in \{1,\dots,n\}$

and $$j = \begin{cases} i & i \le n_\sigma \\ n_\sigma & i > n_\sigma \end{cases} \qquad (8)$$

and $$T = \prod_{p=i}^{n_\sigma - 1} \prod_{q=p+1}^{n_\sigma} T_{pq}(\alpha'_j)$$

With $j = 1/2(2n_\sigma - p)\cdot(p+1) - 2n_\sigma + q$ and the rotation matrices

$$T_{pq}(\alpha'_j) := \begin{pmatrix} 1 & 0 & & & & & & 0 \\ 0 & 1 & & & & & & \\ & & \cos\alpha'_j & & & -\sin\alpha'_j & & \\ & & & 1 & & & & \\ & & & & 1 & & & \\ & & \sin\alpha'_j & & & \cos\alpha'_j & & \\ & & & & & & 1 & \\ & & & & & & 1 & 0 \\ 0 & & & & & & 0 & 1 \end{pmatrix} \qquad (9)$$

which proceeds by first mutating the strategy parameters $\vec{\sigma}$ and $\vec{\alpha}$ and then modifying $\vec{x}$ according to the new set of strategy parameters obtained from mutating $\vec{\sigma}$ and $\vec{\alpha}$:

- $m_\sigma : R_+^{n_\sigma} \to R_+^{n_\sigma}$ mutates the strategy parameters $\vec{\sigma}$:

$$m_\sigma(\vec{\sigma}) = \vec{\sigma}' = (\sigma_1 \exp(z_1 + z_0), \ldots, \sigma_{n_\sigma} \exp(z_{n_\sigma} + z_0)) \qquad (5)$$

where $z_0 \sim N(0, \tau_0^2)$, $z_i \sim N(0, \tau^2)$ $\forall i \in \{1, \ldots, n_\sigma\}$.

Again lower bound ε_σ is enforced for all σ_i.

- $m_\alpha : R^{n_\alpha} \to R^{n_\alpha}$ mutates $\vec{\alpha}$:

$$m_\alpha(\vec{\alpha}) = \vec{\alpha}' = (\sigma_1 + z_1, \ldots, \sigma_{n_\alpha} + z_{n_\alpha}) \qquad (6)$$

where $z_i \sim N(0, \beta^2)$ $\forall i \in \{1, \ldots, n_\alpha\}$

Rotating angles are kept feasible (i.e., in the interval $[-\pi, \pi]$) by circularly mapping them into the feasible range whenever it is left by mutation.

- $m_x : R^n \to R^n$ mutates the object variables $\vec{x}$, using the already mutated $\vec{\sigma}'$ and $\vec{\alpha}'$:

$$m_x(\vec{x}) = \vec{x}' = (x_1 + cor_1(\vec{\sigma}', \vec{\alpha}'), \ldots, x_n + cor_n(\vec{\sigma}', \vec{\alpha}')) \qquad (7)$$

α_{12} represents the rotation angle of the hyperellipsoid. In the general case of correlated mutation, hyperellipsoid may align itself arbitrarily in the n-dimensional search space [2][6].

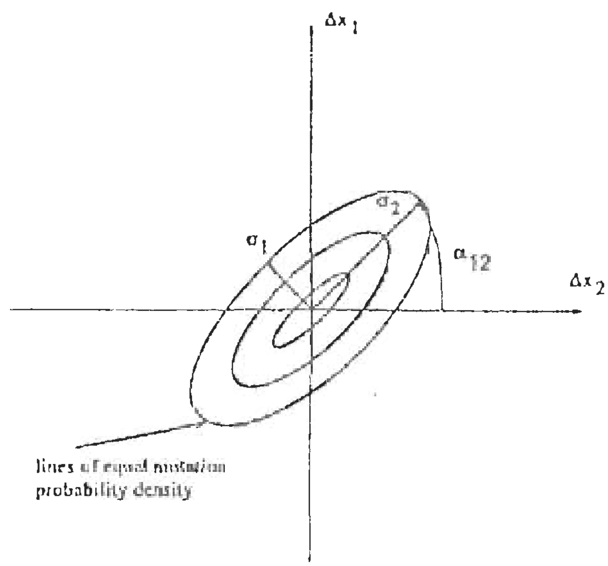

Figure 1 Correlated Mutation for case $n = 2, n_\sigma = 2$ and $n_\alpha = 1$.

Then, the correlated mutation operator $m_{\{\tau_0,\tau,\beta\}} : I \to I$ is defined as follows[2] [2]:

$$m_{\{\tau_0,\tau,\beta\}}(\vec{a}) = m_x(\vec{x}) \circ (m_\sigma(\vec{\sigma}) \times m_\alpha(\vec{\alpha})) = (\vec{x}', \vec{\sigma}', \vec{\alpha}') \qquad (4)$$

[2] where × stands for Cartesian product relation.

- $m_\sigma : R_+^{n\sigma} \to R_+^{n\sigma}$ **mutates the strategy parameters** $\vec{\sigma}$ **:**

$$m_\sigma(\vec{\sigma}) = \vec{\sigma}' = (\sigma_1 \exp(z_1 + z_0), \ldots, \sigma_{n_\sigma} \exp(z_{n_\sigma} + z_0)) \quad (2)$$

where $z_0 \sim N(0, \tau_0^2)$, $z_i \sim N(0, \tau^2)$ $\forall i \in \{1, \ldots, n_\sigma\}$ **.**

To prevent standard deviations from becoming practically zero, a minimal value of ε_σ **is algorithmically enforced for all** σ_i **.**

- $m_x : R^n \to R^n$ **mutates the object variables** $\vec{x}$ **:**

$$m_x(\vec{x}) = \vec{x}' = (x_1 + z_1, \ldots, x_n + z_n)$$

where **(3)**

$$z_i \sim N(0, \sigma_j'^2) \quad \forall i \in \{1, \ldots, n\} \quad and \quad j = \begin{cases} i & i < n_\sigma \\ n_\sigma & i > n_\sigma \end{cases}$$

2. Correlated Mutation in Evolution strategies

Then, Schwefel follows [1][4][5] proposed the concept of correlated mutations by including additional strategy parameters $\vec{\alpha}$ **, which represent the angles of the principal axes of the generalized n-dimensional normal distribution.**

The basic idea of correlated mutation is captured here for the case $n = 2$**,** $n_\sigma = 2$, **and** $n_\alpha = 1$ **in figure 1, where the lines of equal mutation probability density of the two-dimensional normal distribution are plotted. Notice that the standard deviations** σ_1 **and** σ_2 **determine the relation of the length of the main axes of the hyperellipsoid, and**

Table 1: Notational conventions used

Notation	Description
I	Space of individuals
$\vec{a} \in I$	A single individual
$\vec{x} \in R^n$	Vector of object variables
$f : R^n \to R$	Objective function
$\Phi : I \to R$	Fitness function
μ	Parent population size
λ	Offspring population size
$P(t) = \{\vec{a}_1(t),...,\vec{a}_\mu(t)\}$	Population at generation t
$m_{\Theta_m} : I \to I$	Mutation operator (with parameter set Θ_m)

Un correlated mutation operator consists of the addition of a normally distributed random number to each component of the object variable vector, corresponding to a step in the search space. The variance of the step-size distribution is itself subject to mutation as a strategy variable. That is $m_{\{\tau_0,\tau\}} : I \to I$, is defined as follows [1] [3]:

$$m_{\{\tau_0,\tau\}}(\vec{a}) = m_x(\vec{x}) \circ m_\sigma(\vec{\sigma}) = (\vec{x}', \vec{\sigma}') \qquad (1)$$

Which proceeds by first mutating the strategy parameters $\vec{\sigma}$ and then modifies $\vec{x}$ according to the new set of strategy parameters obtained from mutating $\vec{\sigma}$:

[1] Where $\circ$ stands for composition relation.

1. Introduction

In contrast to exogenously rule for changing the standard deviations of mutations (i.e., the 1/5 success rule) in (1+1)-ES, Schwefel takes a more general method that is inspired from nature to adjust the mutation parameters (internal model) of the algorithm. A method was found by taking a closer look at the natural model, where the genotype itself incorporates mechanisms to control its own mutability (by means of genotype segments that encode repair enzymes, or by so-called mutator genes). Transferring this to the evolution strategy means that the standard deviation for mutation must evolves by means of mutation and recombination just as the object variables do in a process called self-adaptation of the strategy parameters [1][2].

According to the generalized structure of $(\mu \overset{+}{,} \lambda)$-ESs individuals, there are two forms of mutation: uncorrelated and correlated mutation operators. This section reviews the rule of uncorrelated mutation. Section 2, then presents in some details the correlated mutation as will be used in what follows. Section 3 then presents Rosenbrock's Saddle problem, its difficulties and the previous attempts to track this problem. Finally section 4 presents evolution strategies with our proposed setting for strategy parameters $\vec{\sigma}$ and $\vec{\alpha}$, their length n_σ and n_α, type of perturbation operators and the experimental results. Conclusions are drawn in section 5.

Before continue, table 1 presents some notational conventions used in what follow.

An Evolution Strategy With Correlated Mutation For Solving Rosenbrock's Saddle Problem

Dr. Mayada F. Abdul-Halim*

Abstract

The Rosenbrock's Saddle problem is a unimodal optimization problem that has the minimum optimum located within a long narrow curved valley which is only slightly decreasing and with a flat bottom. This valley is located near the boundaries of the search space. To find the valley is trivial; however, convergence to the global minimum is difficult because the problem has a strong inter-parameter linkage or non-separable characteristic. The multimembered evolution strategies, $(\mu \overset{+}{,} \lambda)$-ESs can be used as search algorithm to find this minimum optimum. This paper presents an evolution strategy with appropriate setting for their operators and strategy parameters that enable the algorithm to solve Rosenbrock's Saddle problem reliably.

* University of Baghdad - College of Science - Department of Computer Science